National Cancer Institute

measurementERRORwebinarseries

Accounting for measurement error in dietary intake data

September 20-December 6, 2011

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

Presenters and Collaborators

Sharon Kirkpatrick
Series Organizer

Regan Bailey	Laurence Freedman	Douglas Midthune
Dennis Buckman	Patricia Guenther	Amy Subar
Raymond Carroll	Victor Kipnis	Fran Thompson
Kevin Dodd	Susan Krebs-Smith	Janet Tooze

measurementERRORwebinarseries

2

---

measurementERRORwebinarseries

This series is dedicated to the memory of
Dr. Arthur Schatzkin

In recognition of his internationally renowned contributions to the field of nutrition epidemiology and his commitment to understanding measurement error associated with dietary assessment.

Webinar series goal

- Provide participants with an understanding of:
  - The sources and magnitudes of dietary measurement errors
  - How measurement error may affect estimates of usual dietary intake distributions
  - How measurement error may affect analyses of diet-health relationships
  - How the effects of measurement error may be mitigated

measurementERRORwebinarseries

4

National Cancer Institute

measurementERRORwebinarseries

Introduction to measurement error in dietary intake data

Sharon Kirkpatrick
National Cancer Institute

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

Today's objective

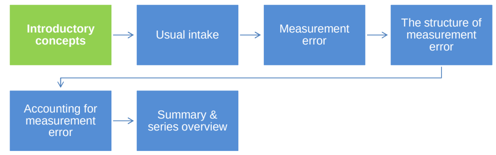
- Participants will gain an understanding of:
  - The concept of usual dietary intake
  - Sources of measurement errors and their impact
  - Concepts underpinning approaches to reducing and correcting for measurement error

measurementERRORwebinarseries

6 Introduction to measurement error in dietary intake data

Outline

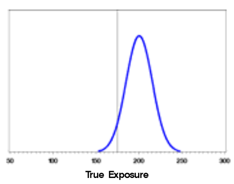
- Introductory concepts
- Usual intake
- Measurement error
- The structure of measurement error
- Accounting for measurement error
- Summary & series overview

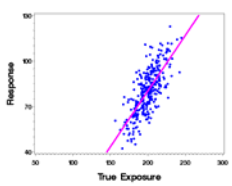


INTRODUCTORY CONCEPTS

Two main areas of interest

- **Describing usual intake distributions**: mean, percentiles, proportion above or below a threshold





- **Estimating diet-health relationships**: regression coefficients

Two main areas of interest - caveat

- For estimating usual intake distributions or diet-health relationships, draw upon data for a population/group of interest rather than a standalone individual
- In contrast to clinical settings, where interest is in a standalone individual and only his or her data

Two types of self-report instruments

- **Short-term instruments** *(e.g., 24-hour recalls, food records, food diaries)*
  - Often used in population surveys for monitoring health and nutrition
- **Long-term instruments** *(e.g., food frequency questionnaire)*
  - Often used in large cohort or case-control studies to examine diet-health relationships

Main versus reference instrument

- **Main instrument**
  - The primary dietary assessment instrument
- **Reference instrument**
  - An instrument used to calibrate or validate the main instrument
  - Assumed to provide estimates that are closer to the underlying truth than the main instrument (alloyed gold standard)

Instruments drawn upon in this series

- **24-hour recall (24HR)**
  - Main instrument for estimating usual intake distributions
  - Reference instrument for estimation of diet-health relationships using food frequency questionnaire as main instrument
  - For future studies, main instrument for assessing diet-health relationships

Instruments drawn upon in this series

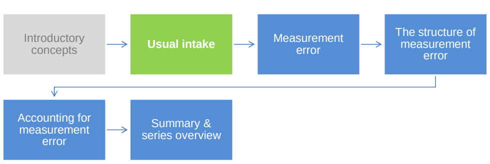
- **Food frequency questionnaire (FFQ)**
  - Main instrument for assessing diet-health relationships
  - Supplemental instrument for studying diet-health relationships using a short-term instrument as the main instrument

Instruments drawn upon in this series

- **Recovery biomarker (reference instrument)**
  - Specific biologic product that is directly related to intake and not subject to homeostasis or substantial interindividual differences in metabolism
  - Examples:
    - Doubly labeled water for energy intake
    - Urinary nitrogen for protein intake

Daily vs. episodic consumption

- **Consumed nearly daily by nearly all persons**
  - E.g., vitamin C, total grains, total vegetables, solid fats, added sugars
- **Consumed episodically by most persons**
  - E.g., vitamin A, whole grains, dark green vegetables, fish



USUAL INTAKE

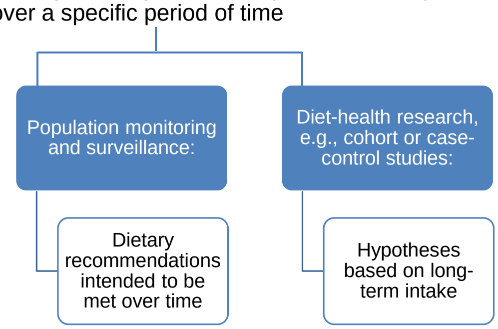
Usual dietary intake

Average or long-run intake (*habitual intake*) over a specific period of time

Usual intake

Usual dietary intake

Average or long-run intake (*habitual intake*) over a specific period of time



measurementERRORwebinarseries

19 Introduction to measurement error in dietary intake data

Usual intake

Challenges to estimating usual intake

- Not directly observable
- Typically rely on self-report instruments
  - Measure usual intake with error

measurementERRORwebinarseries

20 Introduction to measurement error in dietary intake data

Usual intake

Food frequency questionnaire (FFQ)

measurementERRORwebinarseries

21 Introduction to measurement error in dietary intake data

Usual intake

Food frequency questionnaire (FFQ)

- Aims to capture long-term intake
- Cognitively challenging
- Affected by recent diet
- Finite food list
- Lack of detail → assumptions required in converting to nutrient and food group intake

measurementERRORwebinarseries

22 Introduction to measurement error in dietary intake data

Usual intake

Food frequency questionnaire (FFQ)

- Aims to capture long-term intake
- *Inexpensive to administer*
- Cognitively challenging
- Affected by recent diet
- Finite food list
- Lack of detail → assumptions required in converting to nutrient and food group intake

measurementERRORwebinarseries

23 Introduction to measurement error in dietary intake data

Usual intake

24-hour recall (24HR)

measurementERRORwebinarseries

24 Introduction to measurement error in dietary intake data

Usual intake

24-hour recall (24HR)

- Less cognitively challenging (relies on short- term recall)
- Rich detail → fewer assumptions required in converting to nutrient and food group intake
- Aims to capture recent diet
  - Need more than one to assess usual intake

measurementERRORwebinar series

25 Introduction to measurement error in dietary intake data

Usual intake

24-hour recall (24HR)

- Less cognitively challenging (relies on short- term recall)
- Rich detail → fewer assumptions required in converting to nutrient and food group intake
- Aims to capture recent diet
  - Need more than one to assess usual intake
- *Expensive to collect and code (until recently)*

measurementERRORwebinar series

26 Introduction to measurement error in dietary intake data

Usual intake

Food records/diaries

measurementERRORwebinar series

27 Introduction to measurement error in dietary intake data

Usual intake

Food records/diaries

- Less cognitively challenging (does not rely on memory)
- Rich detail → fewer assumptions required in converting to nutrient and food group intake
- Aims to capture current diet (often over several consecutive days)
- Recording may affect intake (reactivity)

measurementERRORwebinar series

28 Introduction to measurement error in dietary intake data

Usual intake

Food records/diaries

- Less cognitively challenging (does not rely on memory)
- Rich detail → fewer assumptions required in converting to nutrient and food group intake
- Aims to capture current diet (often over several consecutive days)
- Recording may affect intake (reactivity)
- *Expensive to code (until recently)*

measurementERRORwebinar series

29 Introduction to measurement error in dietary intake data

Usual intake

Challenges to estimating usual intake

- Self-report instruments used to assess usual dietary intake are affected by several types of measurement error
  - If we ignore this error, our results may be biased

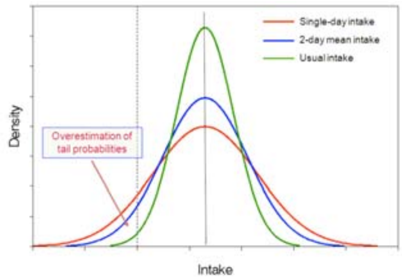
measurementERRORwebinar series

30 Introduction to measurement error in dietary intake data

Usual intake

Describing intake distributions

- Using observed rather than true intake can lead to erroneous conclusions

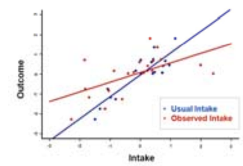


measurementERRORwebinarseries
31 Introduction to measurement error in dietary intake data

Usual intake

Estimating diet-health relationships

- Regression using observed rather than true intake produces:
  - Attenuated slope estimate



  - Loss of power to detect relationship between a dietary exposure and an outcome

measurementERRORwebinarseries
32 Introduction to measurement error in dietary intake data

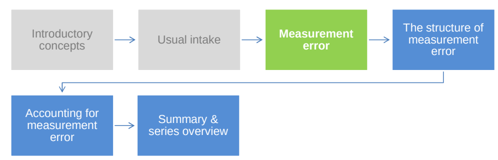
Usual intake

Challenges to estimating usual intake

- Self-report instruments used to assess usual dietary intake are affected by several types of measurement error

Need to understand and address error to avoid biased results

measurementERRORwebinarseries
33 Introduction to measurement error in dietary intake data



MEASUREMENT ERROR

measurementERRORwebinarseries
34 Introduction to measurement error in dietary intake data

Measurement error

What is measurement error?

- Difference between the true value and the value obtained from a measure

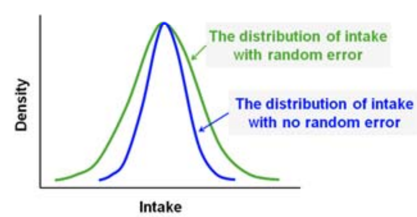
Random | Systematic

measurementERRORwebinarseries
35 Introduction to measurement error in dietary intake data

Measurement error

Random error

- An unpredictable source of error that contributes variability
  - Instrument may be accurate (i.e., unbiased), but may not be precise



measurementERRORwebinarseries
36 Introduction to measurement error in dietary intake data

Measurement error

Random error

- An unpredictable source of error that contributes variability
  - Instrument may be accurate (i.e., unbiased), but may not be precise

If an instrument has only random error, the average of many repeat measures approximates the true value

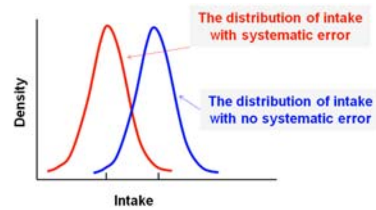
measurementERRORwebinarseries

37 Introduction to measurement error in dietary intake data

Measurement error

Systematic error (bias)

- A source of error in which measurements consistently depart from the true value in the same direction
  - Instrument is inaccurate (i.e., biased*)*



measurementERRORwebinarseries

38 Introduction to measurement error in dietary intake data

Measurement error

Systematic error (bias)

- A source of error in which measurements consistently depart from the true value in the same direction
  - Instrument is inaccurate (i.e., biased*)*

If an instrument has systematic error, the average of many repeat measures *does not* approximate the true value

measurementERRORwebinarseries

39 Introduction to measurement error in dietary intake data

Measurement error

Key measurement error terms

- **Random error:**
  - Within-person random error
- **Systematic error:**
  - Person-specific bias
  - Constant additive error
  - Intake-related bias

measurementERRORwebinarseries

40 Introduction to measurement error in dietary intake data

Measurement error

Key measurement error terms

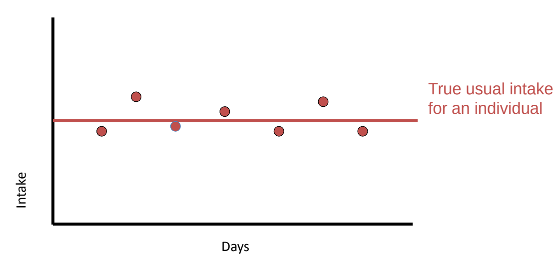
- **Random error:**
  - Within-person random error
- **Systematic error:**
  - Person-specific bias
  - Constant additive error
  - Intake-related bias

measurementERRORwebinarseries

41 Introduction to measurement error in dietary intake data

Measurement error

Within-person random error

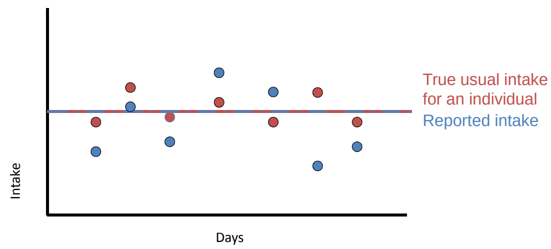


measurementERRORwebinarseries

42 Introduction to measurement error in dietary intake data

Measurement error

Within-person random error



measurementERRORwebinarseries

43 Introduction to measurement error in dietary intake data

Measurement error

Key measurement error terms

- Random error:
  - Within-person random error
- **Systematic error:**
  - Person-specific bias
  - Constant additive error
  - Intake-related bias

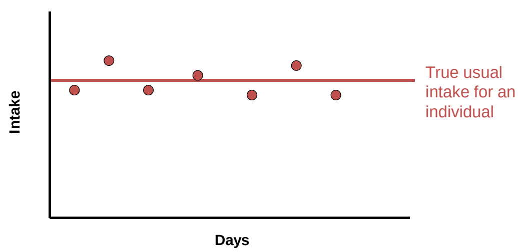
measurementERRORwebinarseries

44 Introduction to measurement error in dietary intake data

Measurement error

Person-specific bias (systematic)

- Defined by personal characteristics



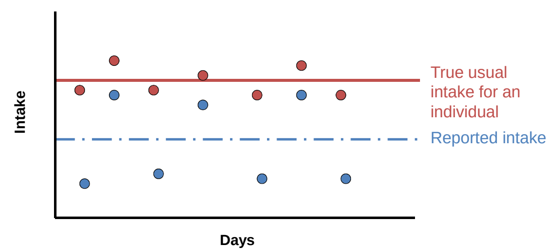
measurementERRORwebinarseries

45 Introduction to measurement error in dietary intake data

Measurement error

Person-specific bias (systematic)

- Defined by personal characteristics



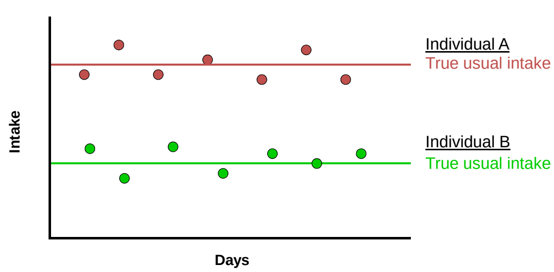
measurementERRORwebinarseries

46 Introduction to measurement error in dietary intake data

Measurement error

Constant additive error (systematic)

- Occurs equally for all participants



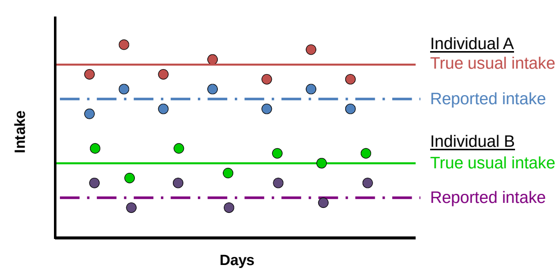
measurementERRORwebinarseries

47 Introduction to measurement error in dietary intake data

Measurement error

Constant additive error (systematic)

- Occurs equally for all participants

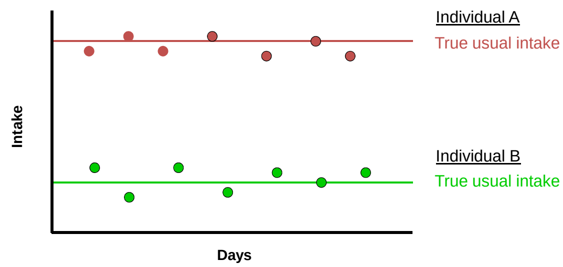


measurementERRORwebinarseries

48 Introduction to measurement error in dietary intake data

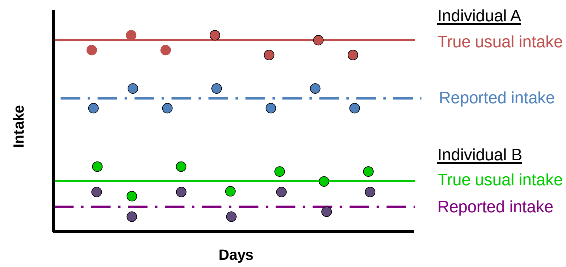
Intake-related bias (systematic)

- Proportional to true intake level



Intake-related bias (systematic)

- Proportional to true intake level



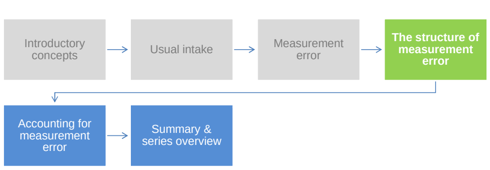
Key measurement error terms

- **Random error:**
  - Within-person random error
- **Systematic error:**
  - Person-specific bias
  - Constant additive error
  - Intake-related bias

Impact of measurement error

- Random error can be dealt with by averaging across repeat measures
- This is not the case for systematic error

The ideal instrument has only random error



THE STRUCTURE OF MEASUREMENT ERROR

The structure of measurement error

$$R_{ij} = \beta_0 + \beta_1 T_i + u_i + e_{ij}$$

Constant additive error (β_0)

Intake-related bias (β_1)

Person-specific bias (u_i)

Random error (e_{ij})

R = reported intake

T = truth

i=person, j=day

How can we study error in a given instrument?

- Validation studies: examine measurement error structure by comparing self-report instrument with a reference instrument
- *Alloyed gold standard* used to estimate truth
  - More extensive self-report instrument
  - Recovery biomarker

Observing Protein and Energy Nutrition (OPEN)

- 261 men and 223 women aged 40-69 years living in Montgomery County, Maryland
- Assessed measurement error structure of:
  - Interviewer-administered 24HR
  - FFQ (Diet History Questionnaire)

OPEN study design

Self-report instruments:	Recovery biomarkers:
• 24HR (2 repeats) • FFQ (2 repeats)	• Doubly labeled water (2 repeats for n=25) • Urinary nitrogen (2 repeats)

OPEN findings: Structure of measurement error

24-hour recall (24HR)	Food frequency questionnaire (FFQ)
• **Larger within-person random error** • Smaller systematic error	• Smaller within-person random error • **Larger systematic error**

Energy underreporting

- Results of OPEN (and other large validation studies) suggest a tendency toward serious energy under-reporting at the group level:
  - 24HR by 10%
  - FFQ by 30%
- Part of systematic error – due to sources of group-level bias

OPEN findings: Attenuation and correlation

- **Attenuation factor**: the degree to which a regression coefficient is biased to the null (attenuated) due to measurement error
  - Closer to zero = more attenuation
- **Correlation between self-report and truth**: related to statistical power to detect diet-health relationships
  - Closer to zero = less powerful the study will be (i.e., need larger sample size)

OPEN findings: Attenuation and correlation

- Attenuation factors and correlation coefficients are substantially better (closer to 1) for repeated 24HR compared to FFQ

OPEN findings: hypothetical scenarios (men)

Nutrient	1 FFQ	2 FFQ	1 24HR	4 24HR	14 24HR
	Attenuation factor/correlation with truth				
Energy	.08/.20	.09/.21	.18/.34	.30/.45	.36/.49
Protein	.16/.32	.17/.34	.20/.37	.37/.51	.46/.57
Protein Density	.40/.43	.49/.47	.23/.38	.50/.55	.68/.65

OPEN findings: hypothetical scenarios (women)

Nutrient	1 FFQ	2 FFQ	1 24HR	4 24HR	14 24HR
	Attenuation factor/correlation with truth				
Energy	.04/.10	.05/.11	.10/.21	.20/.30	.26/.35
Protein	.14/.30	.16/.32	.14/.29	.32/.44	.46/.53
Protein Density	.32/.35	.38/.38	.16/.25	.40/.39	.61/.49

Summary: Measurement error structure

- Main sources of error for dietary components for which we can assess measurement error structure:
  - 24HR: random within-person error
    - Can be mitigated by repeats and averaging

Summary: Measurement error structure

- Main sources of error for dietary components for which we can assess measurement error structure:
  - FFQ: systematic error
    - Unaffected by averaging
    - Unless we have a reference instrument with only random error, cannot correct for systematic error
      - Available for only 2 to 3 dietary components

Summary: Measurement error structure

- Studies using 24HR as main instrument:
  - Intake distributions closer to truth because can account for random error
  - Less bias and more power to detect diet-health relationships

Use the instrument with the smallest systematic error (i.e., 24HR) – assume it is *unbiased*

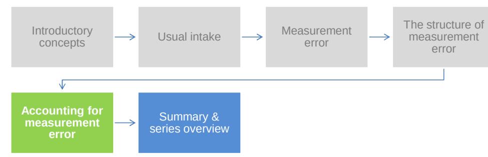
The structure of measurement error

Implications for study design and analysis

- Use of 24HR with repeats for estimation of usual intake distributions
- Efforts to make collection of multiple 24HR from large samples in diet-health studies feasible
  - Development of statistical techniques for the use of short-term instruments (e.g., 24HR) as main instrument in diet-health studies
- Use of 24HR as a reference instrument for studies using FFQ as the main instrument

measurementERRORwebinarseries

67 Introduction to measurement error in dietary intake data



METHODS OF ACCOUNTING FOR MEASUREMENT ERROR

measurementERRORwebinarseries

68 Introduction to measurement error in dietary intake data

Accounting for measurement error

Addressing error in 24HR data

24-hour recall (24HR)	Food frequency questionnaire (FFQ)
• **Larger within-person random error** • Smaller systematic error	• Smaller within-person random error • **Larger systematic error**

- If repeat measures available, can distinguish random from systematic within-person variation and correct for within-person variation

measurementERRORwebinarseries

69 Introduction to measurement error in dietary intake data

Accounting for measurement error

Estimating usual intake distributions

- 24HR with repeats – general approach:
  - Separate within- and between-person variation
  - Estimate distribution of usual intake by removing within-person variation using statistical modeling
    - May also account for nuisance effects (e.g., day of week, recall sequence, interview mode)

Webinars 2, 3, 5

measurementERRORwebinarseries

70 Introduction to measurement error in dietary intake data

Accounting for measurement error

Some existing methods

- U.S. National Research Council (NRC)/Institute of Medicine (IOM)
- Iowa State University (ISU) Method
- U.S. National Cancer Institute (NCI) Method
- EFCOVAL Consortium Multiple Source Method (MSM)
- Statistical Program for Age-adjusted Dietary Assessment (SPADE)

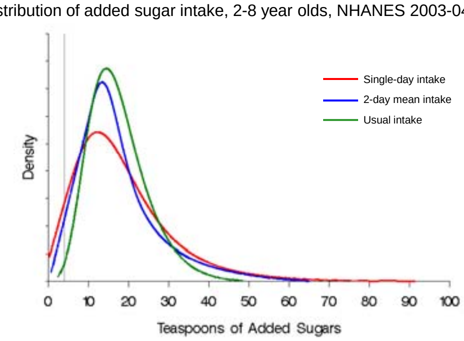
measurementERRORwebinarseries

71 Introduction to measurement error in dietary intake data

Accounting for measurement error

Describing intake distributions

Distribution of added sugar intake, 2-8 year olds, NHANES 2003-04



measurementERRORwebinarseries

72 Introduction to measurement error in dietary intake data

Accounting for measurement error

Estimating group mean intake

- Assumption that 24HR is subject to random error only = unbiased for estimating group mean intake
  - Mean from single 24HR may be sufficient

measurementERRORwebinar series

73 Introduction to measurement error in dietary intake data

Accounting for measurement error

Addressing error in FFQ data

24-hour recall (24HR)	Food frequency questionnaire (FFQ)
• Larger within-person random error • Smaller systematic error	• Smaller within-person random error • **Larger systematic error**

- Use a reference instrument to distinguish the components of systematic error and correct intake estimates

measurementERRORwebinar series

74 Introduction to measurement error in dietary intake data

Accounting for measurement error

Estimating diet-health relationships

- FFQ as the main instrument – general approach:
  - Adjust regression coefficients for bias due to measurement error (regression calibration)
    - Requires data from a reference instrument (e.g., 24HR) administered to a subsample (calibration substudy)

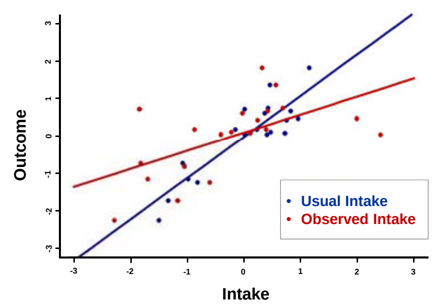
Webinars 6-8

measurementERRORwebinar series

75 Introduction to measurement error in dietary intake data

Accounting for measurement error

Estimating diet-health relationships



measurementERRORwebinar series

76 Introduction to measurement error in dietary intake data

Accounting for measurement error

Estimating diet-health relationships

- Future studies using a 24HR (or other short-term instrument) as the main instrument – general approach:
  - Adjust regression coefficients for bias due to measurement error (regression calibration)
  - Information from FFQ may be used to supplement data from short-term instrument

Webinar 12

measurementERRORwebinar series

77 Introduction to measurement error in dietary intake data

Accounting for measurement error

Alleviating the effects of measurement error

- Combine self-report instruments (e.g., 24HR and FFQ) or self-report and biomarker data to:
  - Improve power to detect relationships
  - Increase precision of estimates, e.g., percentiles of the distribution

Webinars 10 & 11

measurementERRORwebinar series

78 Introduction to measurement error in dietary intake data

Introductory concepts → Usual intake → Measurement error → The structure of measurement error → Accounting for measurement error → **Summary & series overview**

SUMMARY & SERIES OVERVIEW

Today's webinar: summary of main messages

- Self-report instruments used to assess usual dietary intake are affected by several types of measurement error
- Ideal instrument has only random error (unbiased)
- Structure of error in 24HR makes it the best approximation of an unbiased instrument for estimation of intake distributions and diet-health relationships
- Use best instrument possible and statistical techniques to account for measurement error

Today's webinar: the basics

The basics → Describing usual intake distributions → Assessing relationships → Latest developments

- Usual dietary intake
- Random and systematic measurement errors and their impact
- Concepts underpinning approaches to reducing measurement error

Webinars 2-5

The basics → **Describing usual intake distributions** → Assessing relationships → Latest developments

- Describing usual intake distributions for dietary components consumed daily and episodically
- Accounting for complex survey design
- Estimating total usual intake (diet and supplements)

Webinars 6-8

The basics → Describing usual intake distributions → **Assessing relationships** → Latest developments

- Impact of measurement error on assessing diet-health relationships
  - Dietary components consumed daily
  - Episodically consumed dietary components

  } FFQ as main dietary assessment instrument

Webinars 9-12

The basics → Describing usual intake distributions → Assessing relationships → **Latest developments**

- Multivariate applications, e.g., diet quality indices
- Improving estimation by:
  - Combining instruments
  - Combining self-report and biomarker data
- Using 24-hour recalls in diet-health studies

Summary & series overview

Webinar resources

- Series schedule
- Objectives and recommended readings
- Glossary of key terms and notation
- Archived webinars (slides and audio)

riskfactor.cancer.gov/measurementerror

measurementERRORwebinar series

85

Summary & series overview

Other resources

- National Cancer Institute method, references, and SAS macros

  riskfactor.cancer.gov/diet/usualintakes

- NHANES Advanced Dietary Analyses Tutorial

  www.cdc.gov/nchs/tutorials/dietary/advanced

measurementERRORwebinar series

86

QUESTIONS & ANSWERS

Moderator: Kevin Dodd

Please submit questions using the *Chat* function

measurementERRORwebinar series

87 Introduction to measurement error in dietary intake data

National Cancer Institute

measurementERRORwebinar series

Next Session Tuesday, September 27, 2011 10:00-11:30 EDT

Estimating usual intake distributions for foods and nutrients consumed daily by most persons

Kevin Dodd
National Cancer Institute

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES National Institutes of Health